



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
[www.uspto.gov](http://www.uspto.gov)

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/073,427	02/11/2002	Min Chu	M61.12-0397	3970
7590	02/15/2005		EXAMINER	
Steven M. Koehler WESTMAN CHAMPLIN & KELLY International Center, Suite 1600 900 South Second Avenue Minneapolis, MN 55402-3319			SKED, MATTHEW J	
			ART UNIT	PAPER NUMBER
			2655	
DATE MAILED: 02/15/2005				

Please find below and/or attached an Office communication concerning this application or proceeding.

<b>Office Action Summary</b>	<b>Application No.</b>	<b>Applicant(s)</b>	
	10/073,427	CHU ET AL.	
	<b>Examiner</b>	<b>Art Unit</b>	
	Matthew J Sked	2655	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --  
**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

#### Status

- 1) Responsive to communication(s) filed on \_\_\_\_\_.
- 2a) This action is **FINAL**.                    2b) This action is non-final.
- 3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

#### Disposition of Claims

- 4) Claim(s) 1-29 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) Claim(s) \_\_\_\_\_ is/are allowed.
- 6) Claim(s) 1-29 is/are rejected.
- 7) Claim(s) \_\_\_\_\_ is/are objected to.
- 8) Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

#### Application Papers

- 9) The specification is objected to by the Examiner.
- 10) The drawing(s) filed on 11 February 2002 is/are: a) accepted or b) objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

#### Priority under 35 U.S.C. § 119

- 12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) All    b) Some \* c) None of:
  1. Certified copies of the priority documents have been received.
  2. Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
  3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

#### Attachment(s)

- |  |   |
|--|---|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892)  | 4) <input type="checkbox"/> Interview Summary (PTO-413)                     |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948)                                     | Paper No(s)/Mail Date. _____ .  |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)<br>Paper No(s)/Mail Date _____ . | 5) <input type="checkbox"/> Notice of Informal Patent Application (PTO-152) |
|  | 6) <input type="checkbox"/> Other: _____ .                                  |

## **DETAILED ACTION**

### ***Information Disclosure Statement***

1. According to MPEP 37 CFR 1.56, "Each individual associated with the filing and prosecution of a patent application has a duty of candor and good faith in dealing with the Office, which includes a duty to disclose to the Office all information known to that individual to be material to patentability as defined in this section." The published article, "An Objective Measure for Estimating MOS of Synthesized Speech", written by the applicants, outlines several references known prior to filing and are pertinent to the current application. These references are included in the PTO-892 as references cited by the examiner. The applicant is urged to submit any other references of merit on an Information Disclosure Statement for consideration.

### ***Double Patenting***

2. The nonstatutory double patenting rejection is based on a judicially created doctrine grounded in public policy (a policy reflected in the statute) so as to prevent the unjustified or improper timewise extension of the "right to exclude" granted by a patent and to prevent possible harassment by multiple assignees. See *In re Goodman*, 11 F.3d 1046, 29 USPQ2d 2010 (Fed. Cir. 1993); *In re Longi*, 759 F.2d 887, 225 USPQ 645 (Fed. Cir. 1985); *In re Van Ornum*, 686 F.2d 937, 214 USPQ 761 (CCPA 1982); *In re Vogel*, 422 F.2d 438, 164 USPQ 619 (CCPA 1970); and, *In re Thorington*, 418 F.2d 528, 163 USPQ 644 (CCPA 1969).

A timely filed terminal disclaimer in compliance with 37 CFR 1.321(c) may be used to overcome an actual or provisional rejection based on a nonstatutory double patenting ground provided the conflicting application or patent is shown to be commonly owned with this application. See 37 CFR 1.130(b).

Effective January 1, 1994, a registered attorney or agent of record may sign a terminal disclaimer. A terminal disclaimer signed by the assignee must fully comply with 37 CFR 3.73(b).

3. Claims 1-12 are provisionally rejected under the judicially created doctrine of double patenting over claims 1, 9-15, and 18-20 of copending Application No. 10/660,388. This is a provisional double patenting rejection since the conflicting claims have not yet been patented.

The subject matter claimed in the instant application is fully disclosed in the referenced copending application and would be covered by any patent granted on that copending application since the referenced copending application and the instant application are claiming common subject matter, as follows: all the limitations taught in claim 1 of the current application are taught in claim 1 of application 10/660,388 except application 10/660,388 does not recite the limitation "using the relationship to estimate naturalness of synthesized speech". However, the conflicting application does teach estimating the naturalness of the synthesized speech from an objective measure and that the naturalness is a subjective quality, hence the relationship between the objective measure and the subjective measure must be used to estimate the naturalness from the objective measure.

Claims 2-8 of the current application are identical to claims 9-15 of application 10/660,388 and claims 9-11 of the current application are identical to claims 18-20 of application 10/660,388.

Furthermore, there is no apparent reason why applicant would be prevented from presenting claims corresponding to those of the instant application in the other copending application. See *In re Schneller*, 397 F.2d 350, 158 USPQ 210 (CCPA 1968). See also MPEP § 804.

***Claim Objections***

4. Claims 14-16 and 24-27 objected to because of the following informalities:  
“context vectors comprises” should be changed to –context vectors comprise--.  
Appropriate correction is required.

***Claim Rejections - 35 USC § 112***

5. The following is a quotation of the second paragraph of 35 U.S.C. 112:  
*The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter, which the applicant regards as his invention.*
6. Claims 1, 2-9 and 19 are rejected under 35 U.S.C. 112, second paragraph, as failing to set forth the subject matter which applicant(s) regard as their invention. The claims 1 and 19 contradict the invention set forth in the specification. The claims recite the limitation “the objective measure being a function of the textual information derived from the utterances”. However, the specification does not teach any speech to text processing, but instead teaches the opposite as recited in the first paragraph of the summary “The method includes using an objective measure that has components derived from textual information used to form synthesized utterances.” For the purposes of prosecution it will be assumed that the objective measure is a function of the textual information that was used to form the synthesized utterances.
7. The claims 2-9 recite the limitation “the objective measure comprises”, however the objective measure is a mathematical quantity later used to calculate a MOS score, hence it cannot “comprise” a positional or categorical indication for a speech unit. For

the purposes of examination it will be assumed that the "comprises" intends –is computed as a function of--.

***Claim Rejections - 35 USC § 103***

8. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

9. Claims 1, 10, 11 and 19-22 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kitawaki et al. ("Objective Quality Evaluation for Low-Bit-Rate Speech Coding Systems") in view of Rtischev (U.S. Pat. 5,634,086).

As per claim 1, Kitawaki teaches a method for estimating naturalness (quality) of synthesized speech, wherein naturalness is a subjective quality of synthesized speech (a subset of the intelligibility, unintelligible speech not being natural), the method comprising:

generating a set of synthesized utterances (generation of artificial voice, pages 244 and 245);

subjectively rating each of the synthesized utterances (subjectively evaluates the speech using both MOS and Qop, page 243);

calculating a score for each of the synthesized utterances using an objective measure (calculates the LPC cepstrum distance from the cepstrum coefficients, page 242);

ascertaining a relationship between the scores of the objective measure and subjective ratings of the synthesized utterances (compares subjective and objective scores, page 245); and

using the relationship to estimate naturalness of synthesized speech (uses this comparison to determine how well the synthesized speech reflects the characteristics of real speech, page 245).

Kitawaki does not teach the objective measure being a function of textual information that was used to form utterances.

Rtishev teaches a method for calculating the quality of a user's speech by using textual information (number of words in the text, col. 9, line 57 to col. 10 line 9). These references are combinable because they are both related to scoring the quality of a speech signal.

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Kitawaki so the objective measure is a function of textual information that was used to form utterances because the context in which a word is said has an effect on how the word sounds, hence it would give a better measure of the naturalness of the speech to take this into consideration.

10. As per claims 10 and 11, Kitawaki does not teach the objective measure score for each synthesized utterance is a function of a length of said synthesized utterance and this length comprises a number of speech units in an utterance.

Rtishev teaches the objective score is a function of the number of words in the text (col. 9, line 57 to col. 10 line 9).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Kitawaki so the objective measure score for each synthesized utterance is a function of a length of said synthesized utterance and this length comprises a number of speech units in an utterance as taught by Rtischev because the speed which is a function of the length of the utterance has an affect on the sound of the utterance, hence this would give better results.

11. As per claim 19, Kitawaki teaches a method for developing a speech synthesizer, the method comprising:

obtaining a set of synthesized utterances from the speech synthesizer (generation of artificial voice, pages 244 and 245);

subjectively rating naturalness of each of the synthesized utterances (subjectively evaluates the speech using both MOS and Qop, page 243);

calculating a score for each of the synthesized utterances using an objective measure (calculates the LPC cepstrum distance from the cepstrum coefficients, page 242); and

ascertaining a relationship between the scores of the objective measure and ratings of the synthesized utterances (compares subjective and objective scores, page 245).

Kitawaki does not teach the objective measure being a function of textual information of speech units for each of the utterances.

Rtischev teaches a method for calculating the quality of a user's speech by using textual information (number of words in the text, col. 9, line 57 to col. 10 line 9). These

references are combinable because they are both related to scoring the quality of a speech signal.

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Kitawaki so the objective measure is a function of textual information that was used to form utterances as taught by Rtishev because the context in which a word is said has an effect on how the word sounds, hence it would give a better measure of the naturalness of the speech to take this into consideration.

Kitawaki and Rtishev do not teach varying a parameter of the speech synthesizer, obtaining speech units for another utterance after the parameter of the speech synthesizer has been varied, calculating a second score for said another utterance using the objective measure and using the relationship and the second score to estimate naturalness of said another utterance.

However, the Examiner takes Official Notice that recursive training is notoriously well known in the art. Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Kitawaki and Rtishev to vary a parameter of the speech synthesizer, obtain speech units for another utterance after the parameter of the speech synthesizer has been varied, calculate a second score for said another utterance using the objective measure and use the relationship and the second score to estimate naturalness of said another utterance because it would adapt the speech synthesizer for the most natural synthesized speech hence giving better and better results as time progresses.

12. As per claim 20, neither Kitawaki nor Rtischev teach obtaining speech units for another utterance includes obtaining speech units for a second set of utterances, wherein calculating a second score includes calculating corresponding scores for each of the utterances of the second set of utterances, and wherein using the relationship includes using the relationship to estimate naturalness of each of said second set of utterances.

However, the Examiner takes Official Notice that training a speech synthesizer using multiple sets of utterances is notoriously well known in the art. Therefore, the Examiner takes Official Notice that it would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Kitawaki and Rtischev where obtaining speech units for another utterance includes obtaining speech units for a second set of utterances, wherein calculating a second score includes calculating corresponding scores for each of the utterances of the second set of utterances, and wherein using the relationship includes using the relationship to estimate naturalness of each of said second set of utterances because this would allow the synthesizer to produce more natural sounding speech for multiple users hence making the system more pleasurable.

13. As per claim 21, Kitawaki and Rtischev do not teach the parameter comprises an amount of speech units available for synthesis.

However, the Examiner takes Official Notice that excluding speech units prior to synthesis that are unintelligible is notoriously well known in the art. Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to modify the

system of Kitawaki and Rtischev so the parameter comprises an amount of speech units available for synthesis because not synthesizing unintelligible units would give a more natural sounding speech.

14. As per claim 22, Kitawaki and Rtischev do not teach the parameter comprises an algorithm for selecting speech units.

However, the Examiner takes Official Notice that selecting only the speech units that will produce good synthesized speech is notoriously well known in the art. Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Kitawaki and Rtischev by having an algorithm for selecting speech units because this would speed up selection to avoid synthesizing unintelligible speech units so as to give a more natural sounding speech.

15. Claims 2, 3, 6-9, 12-16, and 23-29 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kitawaki in view of Rtischev and in further view of Holm et al. (U.S. Pat. 6,260,016).

16. As per claims 2 and 3, Kitawaki and Rtischev do not teach the objective measure comprises an indication of a position of a speech unit in a phrase or word.

Holm teaches a method for speech to text processing that uses the position of phones within the words as part of the template for synthesis (col. 8, lines 47-48), hence suggesting that the position of a phoneme in a word effects the naturalness of the sound of the phoneme.

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Kitawaki and Rtischev to have objective measure comprise an indication of a position of a speech unit in a phrase or word as suggested by Holm because the analyzed speech unit would be spoken differently according to its relation to the speech units surrounding it hence giving a measure of the speech.

17. As per claims 6 and 7, Kitawaki and Rtischev do not teach the objective measure comprises an indication of a category for the tone of a preceding or following speech unit.

Holm teaches a method for speech to text processing that uses the stress of a syllable that contains multiple phones, hence indicating the stress level of each of the phones within the syllable, as part of the template for synthesis (col. 8, lines 44-47), hence suggesting the tone of surrounding words would affect the naturalness of the target word that is synthesized.

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Kitawaki and Rtischev so the objective measure comprises an indication of a category for the tone of a preceding or following speech unit as suggested by Holm because it would give a better indication of the naturalness of the synthesized speech.

18. As per claim 9, Kitawaki and Rtischev do not teach the objective measure comprises an indication of level of stress of a speech unit.

Holm teaches a method for speech to text processing that uses the stress of the current syllable as part of the template for synthesis (col. 8, line 46), hence suggesting

the stress of the current speech unit would affect the naturalness of the synthesized speech.

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Kitawaki and Rtischev so the objective measure comprises an indication of level of stress of a speech unit as suggested by Holm because the stress placed on a speech unit would affect how it would sound.

19. As per claims 12-16 and 23-27, neither Kitawaki nor Rtischev teach that calculating a score includes generating context vectors for each synthesized utterance wherein the context vectors comprise at least two coordinates of textual information.

Holm teaches determining how well the prosody template corresponds to natural sounding intonation by measuring the area difference between two vectors which represent the duration and mean of the durations of the phones, hence they are context vectors (col. 5, lines 3-25).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Kitawaki and Rtischev to calculate a score including generating context vectors for each synthesized utterance wherein the context vectors comprise at least two coordinates of textual information as suggested by Holm because as taught by Holm this measure would give a good indication of how similar or different samples are from one another (col. 5, lines 5-8).

The coordinates comprising the textual information, which would be used to create the context vectors, would come from the objective measure and the reasons for rejection can be found in regards to claims 2-7 and 9.

Art Unit: 2655

20. As per claims 17 and 28, Kitawaki, Rtishev and Holm do not teach the objective measure includes an indication of a prosodic mismatch between successive speech units.

However, the Examiner takes Official Notice that checking for errors prior to synthesis is notoriously well known in the art. Therefore it would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Kitawaki, Rtishev and Holm to have the objective measure comprise an indication of a prosodic mismatch between successive speech units because an error in prosody matching would create further errors in the system, hence affecting the naturalness of the synthesized voice.

21. As per claims 18 and 29, neither Kitawaki nor Rtishev teach the coordinates are weighted.

Holm teaches the measures can be weighted (col. 5, lines 8-10).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Kitawaki and Rtishev to weight the coordinates as taught by Holm because as Holm states, it would take into account the phyo-acoustic properties (col. 5, lines 8-10).

22. Claims 4 and 5 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kitawaki in view of Rtishev as applied to claim 1 and in further view of Salmi et al. (U.S. Pat. 5,903,655).

As per claims 4 and 5, neither Kitawaki nor Rtischev teach the objective measure comprising an indication of a category for a phoneme preceding or following a speech unit.

Salmi teaches consonants provide more distinguishing information than vowels and that individuals have difficulty understanding speech as being natural if the VC and CV transitions are not properly heard (col. 2, lines 4-12).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Kitawaki and Rtischev to use the category of the phonemes surrounding the speech unit as a measure of naturalness as taught by Salmi because the order of transitions between consonants and vowels affect the masking property of vowels and thus *naturalness* of speech.

23. Claim 8 is rejected under 35 U.S.C. 103(a) as being unpatentable over Kitawaki in view of Rtischev and in further view of Guerra (U.S. Pat. Pub. 2002/0173961A1).

As per claim 8, neither Kitawaki nor Rtischev teach the objective measure comprises an indication of a prosodic mismatch between successive speech units.

Guerra teaches a cause for intelligibility in synthesis is the incorrectness in the prosody (paragraph 3).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Kitawaki and Rtischev to use an indication of a prosodic mismatch between successive speech units as a measure of naturalness as

taught by Guerra because as taught by Guerra when the prosody is incorrect the speech will be difficult or impossible to understand (paragraph 3).

### ***Conclusion***

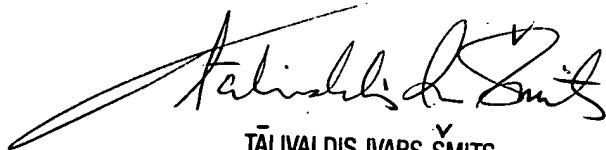
24. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. Chu et al. ("An Objective Measure for Estimating MOS of Synthesized Speech"), Ghitza et al. (U.S. Pat. 6,609,092), Wang et al. ("An Objective Measure for Predicting Subjective Quality of Speech Coders"), Dimolitsas ("Objective Speech Distortion Measures and their Relevance to Speech Quality Assessments"), Cotanis ("Speech Quality Evaluation for Mobile Networks"), Thorpe et al. ("Performance of Current Perceptual Objective Speech Quality Measures"), Bayya et al. (U.S. Pat. 6,446,038), Beerends (U.S. Pat. 6,594,307), Bayya et al. ("Objective Measures for Speech Quality Assessment in Wireless Communications") and Kitawaki et al. ("Quality Assessment of Speech Coding and Speech Synthesis Systems") teach methods for estimating objective measures of the quality of speech. Kochanski et al. (U.S. Pat. 6,810,378) teaches a method for adapting a text-to-speech processor based upon textual information.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Matthew J Sked whose telephone number is (703) 305-8663. The examiner can normally be reached on Mon-Fri (8:00 am - 4:30 pm).

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Talivaldis Smits can be reached on (703) 306-3011. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

MS  
02/08/05



TALIVALDIS IVARS SMITS  
PRIMARY EXAMINER